R E P O R T R E S U M E S

ED 014 759 CG 000 884

EVALUATION ISSUES IN PROGRAMS FOR DISADVANTAGED CHILDREN, THE USE OF NON-TEST PROCEDURES.

BY- MACLENNAN, BERYCE W. LEVINE, MYRNA S.

EDRS PRICE MF-$0.25 HC-$0.56 12P.

DESCRIPTORS- *DISADVANTAGED YOUTH, *EVALUATION, *EDUCATIONAL RESEARCH, *PROGRAM EFFECTIVENESS, EVALUATION METHODS, ENVIRONMENT, PROGRAMS, CULTURALLY DISADVANTAGED, *COUNSELING PROGRAMS,

THE DIMENSIONS OF PROGRAM EVALUATION ARE OUTLINED, SOME NON-TEST METHODS WHICH CAN BE USED ARE DESCRIBED, AND THE IMPORTANCE OF MONITORING AND CONTROLLING THE INTERVENTION ARE STRESSED. THE IMPORTANCE OF TEAMWORK BETWEEN EVALUATORS AND INTERVENTIONISTS IS ILLUSTRATED AND THE VALUE OF SUCH AN APPROACH IN FURTHERING THE CLEAR CONCEPTUALIZATION AND REFINEMENT OF PROGRAMS IS DISCUSSED. SOME PROBLEMS RELEVANT TO PROGRAMS FOR DISADVANTAGED CHILDREN ARE EXAMINED. THE POTENTIAL CAPACITIES OF DISADVANTAGED CHILDREN ARE NOT REFLECTED BY PAST AND PRESENT PERFORMANCE. IN EVALUATING THE EFFECTIVENESS OF INTERVENTION, THE REACTIONS OF THE ENVIRONMENT AS WELL AS THE RESPONSES OF THE SUBJECT MUST BE EXAMINED. UNLESS FAVORABLE CHANGES IN THE INDIVIDUAL ARE SUPPORTED BY THE ENVIRONMENT, SUCH CHANGES ARE UNLIKELY TO BE MAINTAINED. (WR)

ED014759

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM T
PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPIN
STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATIO
POSITION OR POLICY.

EVALUATION ISSUES IN PROGRAMS FOR DISADVANTAGED CHILDREN THE USE OF NON-TEST PROCEDURES*

Beryce W. MacLennan, Ph. D. and Myrna S. Levine, M. A.**

Whenever we undertake any kind of intervention whether it is teaching, training, counselling, therapy, consultation or the institution of some new administrative procedure, we should be concerned with how it works. We evaluate our activities so that we can discard those innovations which do not live up to our expectation, make sure that programs are going according to plan and refine and improve those methods which show promise. "Who, why, what, where, and how?" are as applicable to evaluation as to journalism. We must know what the conditions are that we plan to change, who are involved, what we plan to do, whether the plans are carried out and what their effect is. Thus, there are three stages in evaluation: the original conditions, the process of intervention and the outcome. (1) In this paper we will briefly review some aspects of all three stages and consider in greater detail the evaluation of the intervention process.

Stages in Evaluation

The Baseline. In order to establish the baseline for any program and its evaluation, it is necessary to know the basic characteristics of the population and their present functioning; the characteristics of the interventionists; the setting of the program and the identification of the need for change.

---

*The thinking of the senior author has been particularly influenced by Henry Meyer, Ph. D., and Wyatt Jones, Ph. D. at the Youth Consultation Service of the Episcopal Diocese of New York City 1952-56, Exall Kimbro, M. D., and Milton Shore, Ph. D. at the Mental Health Study Center, Adelphi, Maryland; and Arthur Pearl, Ph. D. at Center for Youth and Community Studies, Howard University, Washington, D. C.

**Dr. MacLennan is Training Center Director, Grant 64221, Office of Juvenile Delinquency and Youth Development, HEW and Associate Director I/C Training, Center for Youth and Community Studies, Howard University, D. C., and Mrs. Myrna S. Levine is Senior Research Supervisor, Training Division, Center for Youth and Community Studies, Howard University, D. C.

CG 000 884

Intervention Process. In order to understand the process of intervention, we must know how change takes place; what is undertaken; how it is carried out; what the immediate responses and interactions are; whether the program is going according to plan; if changes are made; why they are made and what unanticipated side effects may have occurred.

The Outcome. In evaluating outcome, we must look at changes in the population and the setting. Are they according to plan; if so, do these changes produce the results we anticipated; if not, how not and why not.

The Research Design

In establishing our research design, certain facts have to be kept in mind. Resources are always in short supply. A total evaluation would probably cost about 20 times as much as the program itself. Consequently, research should be looked upon as an evolving process. We must establish the questions which have priority for us now and recognize that one set of questions lead to another and that sometimes there are questions we cannot tackle until we have obtained some prior answers or have tried out new methods.

Secondly, the mere fact of evaluating a program may affect it. It has long been recognized that merely paying attention to a person or a program alters the situation. Introducing new, different, or additional people affect the climate of a program. We must be careful to keep an eye on the effects of evaluation on the program and not to swamp it with tests, questionnaires, checklists, or observers.

Teamwork Between Evaluators and Interventionists

As far as possible, the interventions which are planned and the changes which are expected should be derived from theoretical premises and should be defined in operational, behavioral terms. For instance, if we say that our goal is to help someone become less tense or enable them to perform better, this is a subjective judgement which cannot be impartially assessed. However, if we say that we wish to enable a child to be able to remain absorbed in one task for ten, fifteen, or twenty minutes, rather than moving restlessly from task to task every one or two minutes, we can easily judge whether we have been successful.

The program people who are undertaking the intervention should be responsible for the conceptualizations of their program and the clear definition of their methods or the changes they hope to achieve. They must again be able to define their interventions in behavioral terms. They will usually also choose their own population and setting and who will undertake the intervention. Beyond this, however, the establishment of a baseline and the assessment of outcome may be undertaken by an entirely separate team, even though the program people may be engaged in collecting certain data for the evaluators.

When it comes to the evaluation of the Intervention Process, however, we find it much more dynamic and profitable to consider interventionists and evaluators as partners on the same team.

The Intervention Process

The purposes of studying the process of intervention are three-fold: firstly, to know what goes on in the interaction between the interventionist and the subjects; secondly, to assist in ensuring that the program is carried out according to plan; and thirdly, when responses are found to be

different from those expected, to assist in refining and improving the program. Thus such evaluation becomes the ongoing tool of the interventionists.

Both the evaluators and the interventionists will probably participate in data collection. The interventionist will state clearly what he plans to do and what he expects to happen. He will keep descriptive records, make ratings, fill out checklists while the evaluators will act as observers, analyze tapes, conduct tests which will add to the interventionists' data and check on its reliability.

In order for such a team to function effectively, it is important that the evaluators should be looked upon as collaborators and helpers and not as spies, finks, or judges. This means that they must meet together regularly, that the evaluators must be willing to share their results and be able to feedback information speedily enough for it to be useful to the interventionist and that the interventionist must strive to be undefensive, to look upon their work as an adventure or an exploration in which knowledge about the problems they encounter will enable them to achieve their goals more effectively. Mutual trust has to be developed between the two sides of the team and the evaluators must recognize that to remain impartial in such a role is hard and that at times they may be no more reliable than the interventionists. Particularly, they have to watch their overidentification either with the program or with the subjects of the program. Research observers have been known to turn themselves into supervisors or ardent protectors of the pupil or patient. (2)

When we study the process of intervention and attempt to monitor and control program, we have to make choices about the important variables to be studied. Shall we concentrate on the interventionist, what kind

of impression he makes, what kind of person he seems to be, what he does and how he does it? Shall we study the individual subject, his roles, actions, attitudes? Or what goes on in the class or group, the climate, the interactions, the activities, the changing structure? Or possibly we are interested in the overall setting, do changes occur in the institution, its program, its network of communication, its organization of roles and functions in response to our interventions? Obviously, we cannot tackle all of these at once and so we must make decisions which are congruent with our theoretical approaches and our goals.

If our hypothesis is that when a teacher undertakes certain operations, she will elicit certain responses from the children, we must study both whether she is behaving in the way which was planned and whether the children are responding as anticipated.

In one program we hypothesized that there was a relationship between the underachievement of highly gifted youth and the manner in which the family interacted and that if we could help the members together deal differently with problems, we would find a difference in the youth's school performance. (3) We further hypothesized that we could achieve this change in family functioning through bringing families together in a group for a limited number of sessions in which the group leader would focus the discussion on the way the families dealt with their problems. We, thus, had to know the school performance of the youth before and after treatment and how this compared with untreated youth. We needed to know how the family dealt with problems before and after treatment and we required to ensure that the leader focused the group discussion in the desired way. We, therefore, collected data related to the youth's scores on the Lorge-Throndike to indicate that he was gifted. The school record

showed his performance. We devised a problem-solving task (the Family Classification Test) (4) which we administered to the family before and after treatment and as part of a follow-up study, and we devised a scale for rating the group leader's interventions in terms of their relevance and focus to the task which was used by observers with considerable reliability. (3)

In a second program, we were interested in knowing how much was learnt in a course and whether this learning was related to the frequency of exposure. (5) We found that while it was very hard to tell how much had been learnt because information reproduced in a post-test differed considerably from the information which the students showed they possessed in a pre-test, the information which was reproduced correlated positively with the frequency with which such information was transmitted in the course. We were able to reach this conclusion through taping the course and making a content analysis of the discussion.

In a job training program for socially disadvantaged youth, we hypothesized that lack of capacity to take self and group responsibility was related to irresponsibility and poor performance on the job. We used incidence of absenteeism and lateness and teacher ratings of attentiveness on the job as indices of job responsibility and decided that practice in decision-making around job-related problems would be an appropriate training measure. (6) We thus had to collect on the one hand the statistics on job attendance and the teacher ratings and on the other we had to monitor the program to ensure that the students were discussing job-related issues, were taking part in the discussions and that the group members were showing an increased capacity to reach consensus. A group observer, therefore, collected the number of issues discussed,

abstracted the interactions from tapes, had these rated for degree of consensus by independent raters and charted the spread of participation amongst group members.

In developing practicum for new employment roles, we are interested in whether supervisors are giving the trainees the opportunity to perform and learn the tasks which are outlined on a job description. Consequently, we ask our supervisors and trainees to fill out checklists of the tasks which have been performed and our evaluators take sample observations of trainees on the job to check for the reliability of the lists. If we find that trainees are not being afforded adequate experience, we discuss the problem in joint meetings, which may result in the supervisors reorganizing their practices or in a change in the job description. In a classroom aide training program such quality control has resulted in continued refinement of the job description of the aide.

We were interested in knowing in a counsellor training program whether we were teaching the counsellors what we thought we were teaching and in the designated manner and in whether the counsellors were conducting their counselling groups along the lines which we had planned. We consequently studied the degree and quality of participation and the content in both teaching and counselling groups by means of content analysis of tapes, ratings of sample interactions and spot observations. (7)

In an institutional setting where we are interested in the authority structure, we are studying the way in which program is planned, who is involved, how plans are discussed, how decisions are made and who makes them, through an interaction analysis of planning meetings.

Methods of Data Collection

Baseline data is obviously crucial to the evaluation of change, for if we do not know what we started with, we cannot hope to evaluate the effectiveness of our intervention. We must also know or in some way be able to estimate what would happen if we did not intervene or if we continued to intervene in customary ways. We must decide what the relevant variables are and collect our data which may be concerned with standard sociological facts, present performance which can be observed or tested and attitudes and feelings which may be checked, observed or probed for. Thus, we may use questionnaires, structured or semi-structured interviews, checklists, ratings or assessments by self or others, life histories, or assessments of the way in which simulated or real tasks are performed. In one program where we were interested in estimating the attitudes of trainees to disadvantaged populations and their sensitivity to group operations at the start of a program, we asked them to fill out questionnaires, to take part in a group discussion related to problems of disadvanted youth, and to fill out a questionnaire which described the group operations.

The examination of the interaction between interventionist and subject or group of subjects is particularly crucial and also particularly difficult for it is not the content but also the quality of the transaction which is important and a great deal of communication is transmitted nonverbally by gestures and experience, rather than in words. Although a number of researchers have been studying this process, our instruments are still extremely primitive. Again, choices have to be made. Bales, for instance, has defined and refined a method for analyzing each group intervention in terms of 12 categories. (9) In one of our studies, we

have concentrated on the actions of the interventionist. (3) Berne has been interested in the identification of interchange sequences (transactions) which he terms "games." (10) Stock and Lieberman have tried to isolate "critical issues" in therapy groups which can be examined overtime. Leavy has tried to rate the feeling-tone of the participants in (11) terms such as dominating, bossy, compliant. (12)

In the evaluation of intervention process, in our different programs, we have used sociometric data, diary records, checklists of activities, ratings of behavior, interaction scales concerned with quantity and type of interaction, content analysis of interaction and of critical incidents, issues or decisions, and expressions of attitudes by participants.

In some situations we have used depth studies of typical cases or situations in order to gain greater understanding or for the development of hypotheses.

We have gauged our outcomes through studies of behavior as it related to our predictions. We have designed simulated scenes and tasks, obtained evaluations of teachers, supervisors, administrators and students. Used semi-objective criteria such as grades, promotions, capacity to hold the job and other achievements.

We should now like to discuss very briefly some of the problems which are particularly relevant to programs for disadvantaged children.

First, we have to recognize that past and present performance do not reflect the potential capacities of the children. Achievement and intelligence tests give information about what the child is like now, but little help in predicting what he might be able to achieve. We have also come to realize much more vividly that a person's behavior is conditioned not only by his personal characteristics, but also by the circumstances

to which he is exposed. Thus, a child who may behave and learn well under a competent teacher in a calm and structured classroom may fall to pieces in another class where he becomes for some unpredicatable reason a source of irritation to the teacher or where she allows a confused and chaotic climate to prevail. Schools may have very different capacities to tolerate noise, student disruption or a slow pace of learning; and we all know that "drop-outs" may quite frequently be termed "push-outs." Consequently, past behavior cannot really be used as an accurate assessment of a child's present status unless we also know the degree of stress under which the behavior occurred. We thus find very wide variation of responsiveness in children whose original profiles looked rather similar and must constantly be testing our assumptions and predictions.

Secondly, we have also learnt that even if we obtain favorable changes in the individual, if the environment does not support these new ways or worse still, reinforces the old behaviors, such changes are unlikely to be maintained. For instance, if we arouse the children's desire to learn and they are still subjected to conditions under which they are bound to fail, their eagerness will soon evaporate. More difficult is the problem of helping children who have been idle and delinquent in a delinquent world to adopt and maintain more socially acceptable and productive habits, for this not only means a change in behavior and self-concept, it also means a moving away from friends and even family. Similarly, we have found that we may train personnel to perform competently; but if they are placed in an institution which does not accept the same standards and where staff and supervisors are accustomed to different methods, our trainees may well be

forced to abandon their standards or to leave the system. It is, therefore, extremely important in our evaluation of program to keep track not only of what is fed into the program and the responses of the subjects, but also to be concerned with the reactions of the environment within which the individual has to perform.

Summary

In this paper we have outlined the dimensions of program evaluation, described some of the non-test methods which can be used, and stressed the importance of monitoring and controlling the intervention itself and in making careful choices so that we avoid overload. We have tried to illustrate how important it is for evaluators and interventionists to work as a team and to emphasize the value of such an approach in furthering the clear conceptualization and refinement of the program. We cannot overstate that without such conceptualization, no conclusions about the effectiveness of a program can be drawn and the program itself is not repeatable.

We have also noted that past and present records and achievements are not reliable predictors of the potential of socially disadvantaged children and that in evaluating the effectiveness of interventions, it is important to examine not only the responses of the subjects, but the reactions of the environment in which they have to perform. In Lippitt's terms, we must be concerned not only with the individual personality system, but with the institutional and community systems as well.

References

1. Levine, Myrna S. and Arthur Pearl, Designs for the Evaluation of Training Programs. Washington, D. C.: Center for Youth & Community Studies, Howard University, 1965.

2. Krasner, J., et al., "Observing the Observers," International Journal of Group Psychotherapy, 14:214-217, 1964.

3. Kimbro, Exall, et al., "A Multiple Family Group Approach to Some Problems of Adolescence," presented at American Group Psychotherapy Association Annual Conference, 1965.

4. _____, A Family Classification Test. Washington, D. C.: Mental Health Study Center, NIMH, 1965.

5. Burke, Kay R., Beryce W. MacLennan, et al., Group Counselor Training Program Report: Theory, Practice and Evaluation. Washington, D. C.: Center for Youth & Community Studies, Howard University, April 1965.

6. Center for Youth & Community Studies, Howard University, Community Apprentice Program: Disadvantaged Youth in Human Services. Washington, D. C., 1965.

7. Belton, Slyvia I., Beryce W. MacLennan, Myrna S. Levine, and Wilbur H. Jenkins, Jr., "Group Counseling with Freshmen," presented at AGPA Annual Conference, 1966.

8. MacLennan, Beryce W., Philip Goode, and Melvin L. Davis, Counselor Training in Short-term Juvenile Detention Homes. Washington, D. C.: Center for Youth & Community Studies, Howard University, 1966 (in progress).

9. Bales, Robert F., Interaction Process Analysis. Cambridge, Mass.: Addison-Wesley, 1953.

10. Berne, Eric, Games People Play. Brattleboro, Vt.: Grove Press, 1964.

11. Whitaker, D. S. and M. A. Lieberman, Psychotherapy Through the Group Process. New York: Atherton Press, 1964.

12. Leary, Thomas, The Interpersonal Diagnosis of Personality. New York: Ronald Press, 1957.